Medical Expenditure Panel Survey

HOUSEHOLD COMPONENT
MEDICAL CONDITIONS FILE

General File Structure

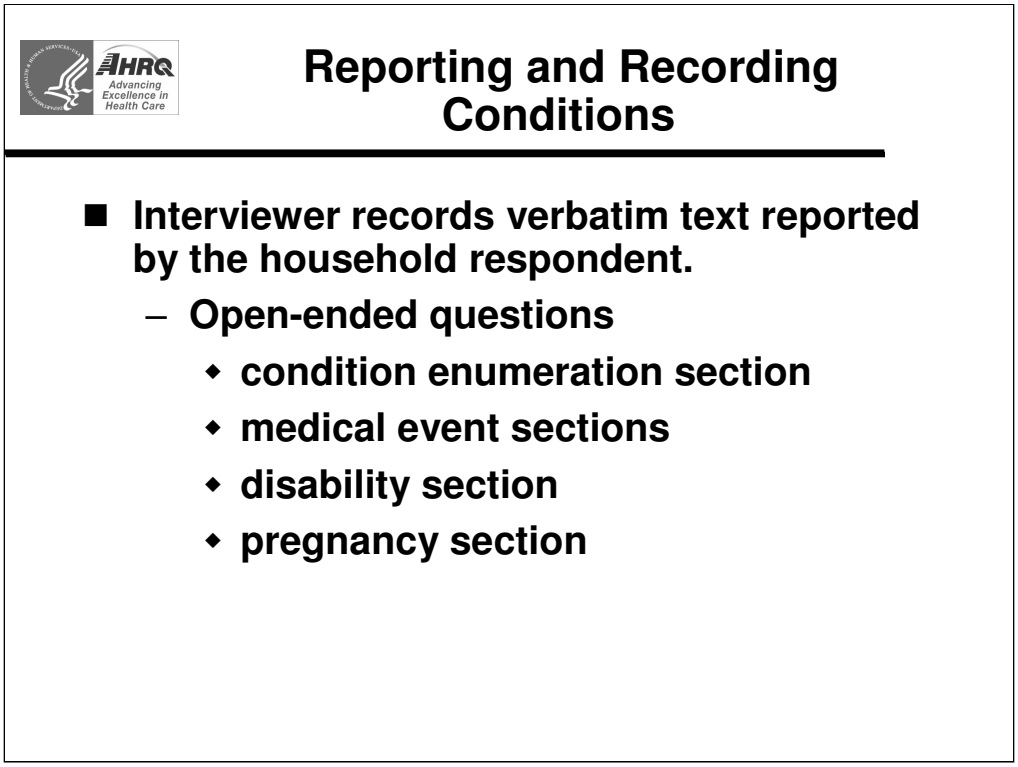
- **Each record represents a unique condition or procedure reported by a household respondent.**
- **Depending on the number of conditions they reported, persons may be represented on the file**
  - **once**
  - **several times**
  - **not at all**

This public use file provides information about the conditions for a nationally representative sample of the civilian noninstitutionalized population of the United States. All information on this file has been collected from household respondents only.

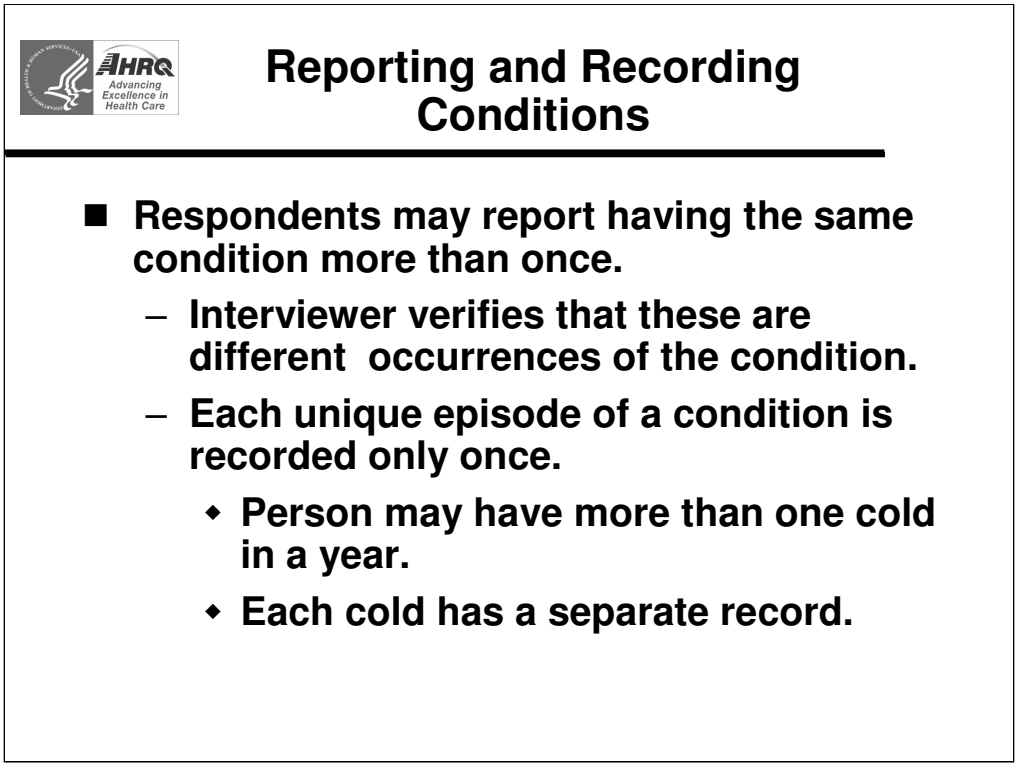
Each record represents a unique condition (defined on the file by an ICD-9 condition code, V-code, or procedure code) reported by a household respondent. If a person reported three different conditions (e.g., asthma, hypertension, and diabetes) then he or she will have three separate records on the condition file. If another respondent reports asthma, bronchitis, and heart disease then he or she will be represented three times, with three separate records on the condition file. Even though each person reported asthma, there will be two separate records for asthma on the file--one for each respondent who reported asthma.

Each record can include information about a condition or a procedure or both.

Although this file contains conditions and procedures it is referred to as “Medical Conditions File.”



Respondents are asked to report current conditions in the condition enumeration section of the questionnaire at every round of data collection. This is an open-ended question and interviewers record the respondent's verbatim text. From this information, CAPI generates a condition roster for every person in the household. Later in the interview, respondents are asked the reason for a medical provider visit, missed workdays, missed school days, and bed-days. At these points in the interview, conditions can be added to the roster if they were not previously mentioned. The only condition recorded in the pregnancy section is the pregnancy itself.



If a respondent reports bronchitis in Round 1 and again in Round 2, the interviewer verifies whether or not this is the same condition that was reported in Round 1. If it is a different condition, then it is entered a second time. If the respondent indicates that it is the same condition reported in Round 1, the interviewer does not enter a new condition. Similarly, if "bronchitis" is reported in Round 1 and later in the interview given as the reason for a provider visit, the interviewer asks if this is the same condition reported previously. If it is a different condition, then it is entered on the condition roster again.

For chronic conditions, such as diabetes or hypertension, there will frequently be several provider visits reported for the same condition, but the condition will appear only once on the person's condition roster and that person will only have one record for that condition on the condition file.

Example of File Structure

DUPERSID	CONDIDX	ICD9CODX	ICD9PROX
1234	12341	asthma	-1
1234	12342	hypertension	-1
1234	12343	diabetes	-1
5432	54321	asthma	-1
5432	54322	cold	chest X-ray
5432	54323	cold	-1
9876	98761	-1	chest X-ray
9876	98762	flu shot	-1

This is an example of what records may look like. The example shows that a person (DUPERSID) can be represented three times on the condition file (three different CONDIDX values), a person may have the same condition (ICD9CODX) more than once, a procedure(ICD9PROX) may be included along with a condition or on a separate record, some records include only a procedure. Flu shot is not a condition, but ICD-9 V-codes (reason for visit) were included on the file if persons gave reason for a provider visit as something other than a condition or a procedure.

Note that the actual ICD9CODX and ICD9PROX codes on the file are numeric values rather than the text used here for illustration. See the next slide and the file documentation.

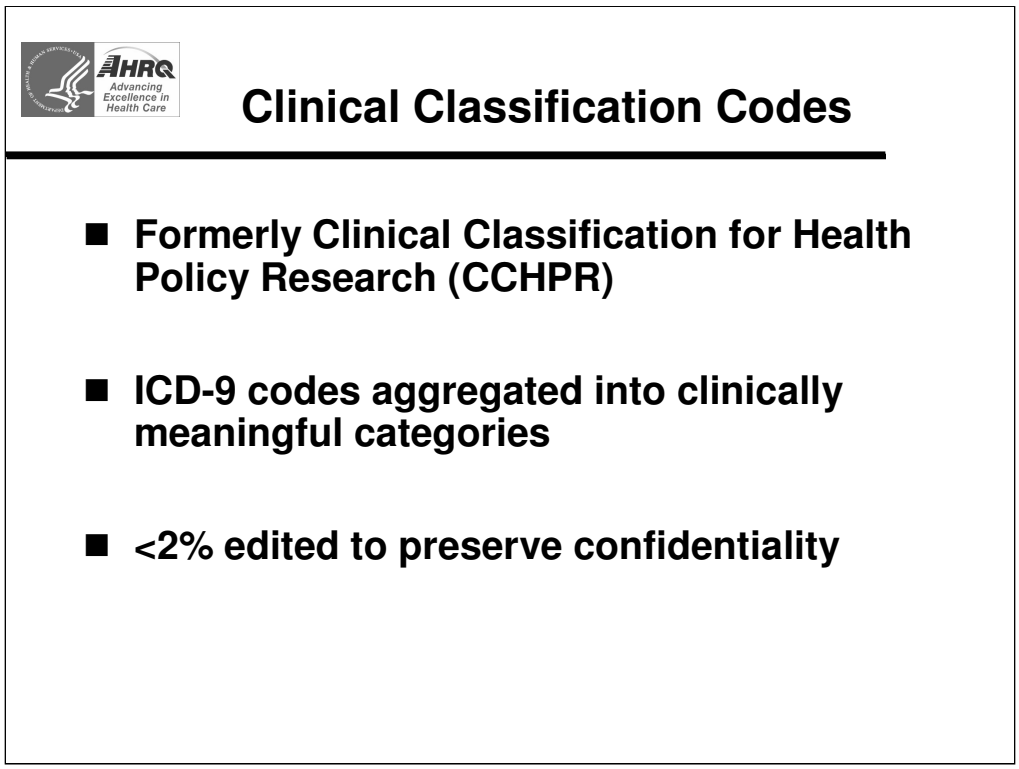
Condition and Procedure Coding

- **Coding**
  - **Professional coders**
  - **Fully specified ICD-9 CM codes (up to 5 digits)**
  - **<2.5% error rates**
- **Editing ICD-9 Codes**
  - **ICD-9 condition codes collapsed to 3 digits to maintain confidentiality**
  - **<10% of condition codes are collapsed further by combining 2 or more 3-digit codes**
  - **Procedure codes collapsed from fully specified (up to 4 digits) to 2-digit codes**
  - **<3% of procedure codes are collapsed further by combining 2 or more 3-digit codes**

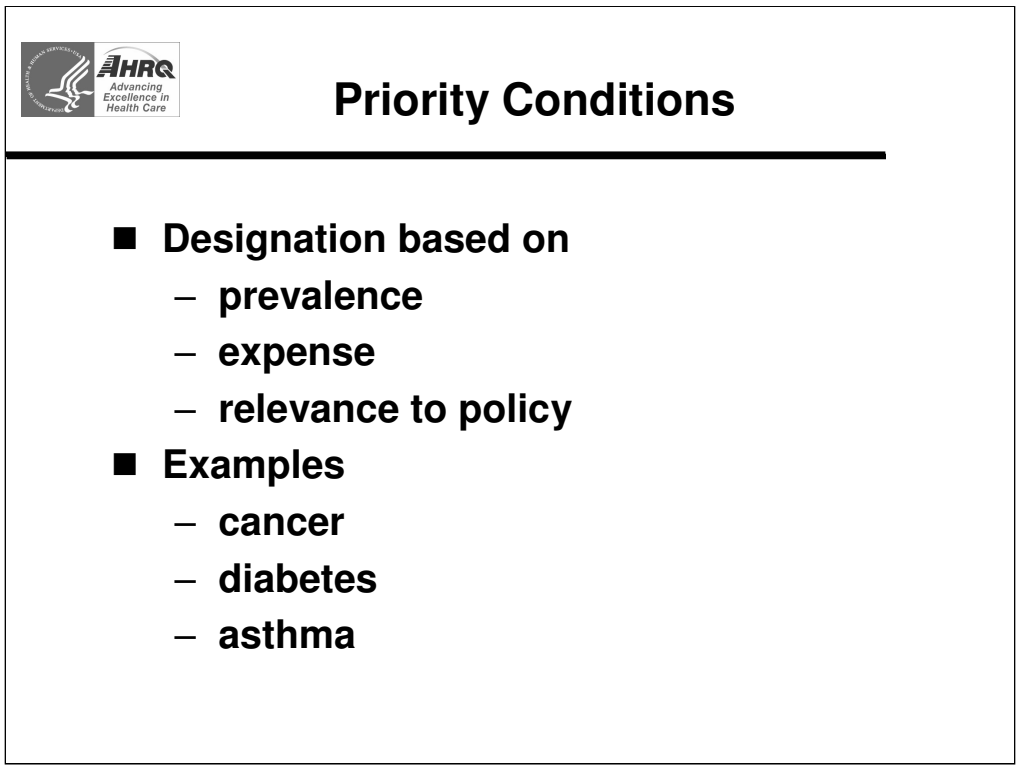
To preserve confidentiality, all of the condition codes provided have been collapsed from fully specified codes to three-digit code categories. Table 1 in Appendix 2 of the documentation provides unweighted and weighted frequencies for all ICD-9 condition code values. To further preserve confidentiality, approximately 10% of the ICD-9 codes were collapsed even further. These are indicated in the codebook in the value label of the ICD-9 code

Procedure codes were also collapsed from fully specified codes to two-digit category codes. To preserve confidentiality approximately 3% were further collapsed. Table 2 in Appendix 2 of the documentation provides unweighted and weighted frequencies for procedures.

Procedures were under reported on the condition roster by household respondents as a reason for a provider visit. Analysts should use procedures identified on the event files for more accurate estimates of procedures.



ICD-9 CM condition codes have been aggregated into clinically meaningful categories that group similar conditions. Clinical Classification Codes aggregate conditions and V-codes into 260 mutually exclusive categories, most of which are clinically homogeneous. Table 3 in Appendix 2 of the documentation provides unweighted and weighted estimates for the Clinical Classification Codes. Appendix 3 lists the ICD-9 codes that have been aggregated for each CCC category.



Certain conditions are designated as "priority" conditions (PRIOLIST = 1). They are identified as such in the field by MEPS interviewers. Some priority conditions are also identified as "injuries." See Appendix 4 of the file documentation for a complete list.

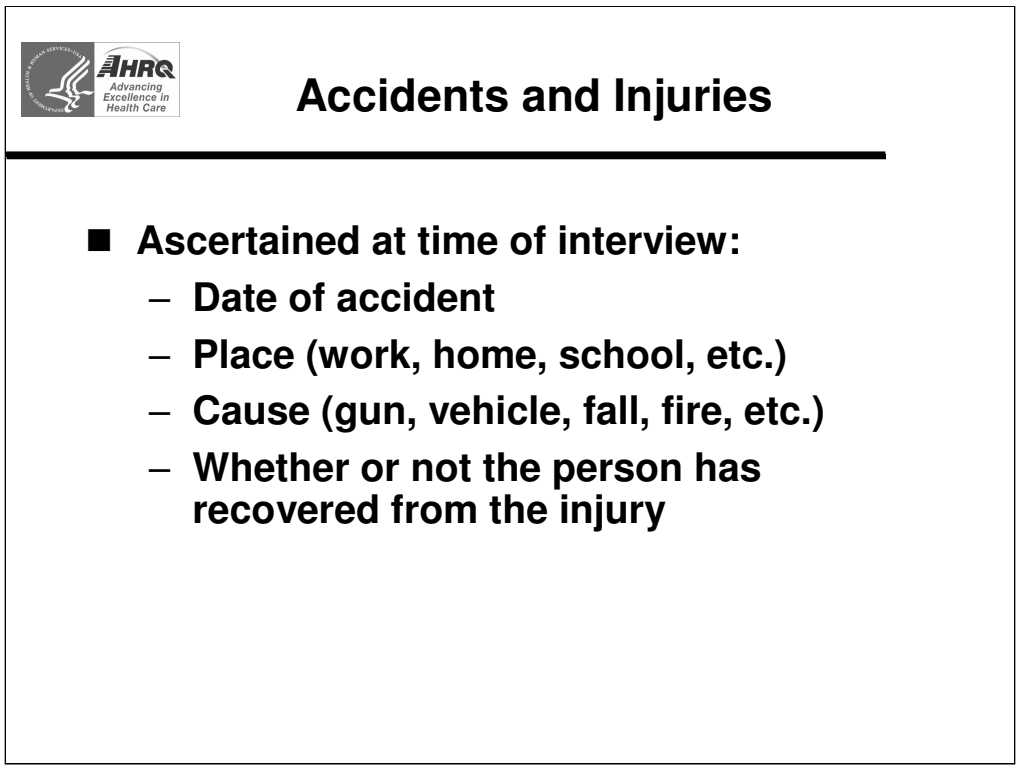
Priority Conditions

- **Specific Priority Condition Questions**
  - **Date condition began**
  - **Round-specific questions**
    - **Was a doctor ever seen or talked to for the condition?**
    - **Did a provider recommend further treatment?**
  - **Follow-Up Questions in Later Round**
    - **Still being treated for this condition?**
    - **How seriously does condition affect overall health and well-being?**

Note that the "date condition began" is not necessarily during the time the person is in the survey. The date a condition began is often many years prior to the survey interview.

Only priority conditions and injuries have date information on the file.

For more details about round-specific variables, see the conditions file documentation.



When a condition was first mentioned, respondents were asked whether it was due to an accident or an injury. If the condition was due to an accident or injury, the interviewer collected the day of the week, month, and year; whether or not the accident or injury occurred at work; where the accident occurred; if it involved a motor vehicle or a gun or other weapon; or if it was the result of poison, fire, drowning, sports, or a fall. Lastly, the interviewer asked if the person had fully recovered from the injury. Because the date of the accident or injury was converted to the day of the week, analysts cannot use the day of the week in conjunction with the accident month and year to determine the exact date.

Only logical edits based on skip patterns were performed for injuries.

Priority and Injury Conditions

- **Date Information**
  - **If a condition is both a priority condition and an injury, date information is on injury date variables.**
- **Previous Year's Information**
  - **For Panel 5 respondents, conditions on the 2001 file first reported in rounds 1 or 2 have round-specific information for those rounds included on the 2000 conditions file.**
  - **PRIORFLG or INJURFLG = 1**

Priority conditions will have date information on the variables CONDBEGD, CONDBEGM, and CONDBEGY (condition begin day, condition begin month, condition begin year).

Injury conditions will have date information on the variables ACCDENTD, ACCDENTM, and ACCDENTY (accident day, accident month, accident year). If a condition is both an injury and a priority condition, only the injury date variables will have data information.

Disability Flag Variables

- Three flag variables
  - YES or NO value (coded 1 or 0)
- Is the condition associated with a
  - Missed work day (MISSWORK)
  - Missed school day (MISSSCHL)
  - Day spent in bed (INBEDFLG)

Due to the MEPS instrument design, there is no link indicating the specific number of disability days associated with a particular medical condition.

Utilization Variables

- **Indicate Number of Events**
  - **ERNUM, HHNUM, IPNUM, OBNUM, OPNUM, RXNUM**
  - **Counts derived from event PUFs**
- **Events for Multiple Conditions**
  - **Events may be associated with more than one condition**
  - **Counts are not person or event level**
  - **Example: One hospital stay for three conditions**
    - **Fractured hip, fractured shoulder, concussion**

Note that DDNUM, a dental utilization variable, is not included on the conditions file beginning in 2001. This is due to a change in the CAPI instrument.

For these three conditions, the conditions file will have three records--each with IPNUM = 1. If you sum these records, you would incorrectly compute three hospital stays for these conditions.

Weighted Estimates

- **Condition file includes person-level weights**
  - **Frequencies from this file will estimate the number of times a condition was reported by the sample population.**
  - **Number of persons reporting a condition can only be estimated at the person level.**

Frequencies derived from the condition-level file will estimate the number of times a condition was reported; e.g., you may want to find the number of head injuries reported in a particular year. If you want to find the number of persons who reported diabetes or asthma, you will need to do this at the person level.

Accuracy of Condition Data

- **Analysts should not presume a high level of precision in condition data.**
  - **Inaccurate or vague reports of condition**
  - **Clustering of ICD-9 codes in NEC (not elsewhere classified)**
  - **One respondent provides information for the entire household**

Although codes were verified and error rates did not exceed 2.5%, analysts should not presume this level of precision in the data; the ability of household respondents to report condition data that can be coded accurately should not be assumed. Reports are sometime vague (e.g., kidney problems, bad knee). This results in clustering of ICD-9 codes as "Not Elsewhere Classified."

Because of these inaccuracies, we do not recommend that analysts use condition data for prevalence or mortality studies.

Comparison of HC and MPC Conditions

- **Substantial lack of congruence between event-specific HC and MPC reports of medical condition:**
  - **Lower reports from HC vs. MPC**
  - **Complex diagnoses have lower agreement**
  - **Comorbidity appears to lower agreement**
  - **Some evidence of social stigma and lack of understanding of medical condition**

Linking to Other MEPS Files

- **ID Variables**
  - **Used to**
    - **identify and distinguish records on a file**
    - **match records in different files**
  - **DUPERSID (person-level ID)**
  - **CONDIDX (condition-level ID)**
  - **EVNTIDX (event-level ID)**
    - **but note PMED differences (LINKIDX, RXRECIDX)**
  - **CLNKIDX (condition-event link ID)**

The MEPS PUF files are constructed so that each record (or, row) contains ID variables that allow you to associate the information on that row with a person or an event.

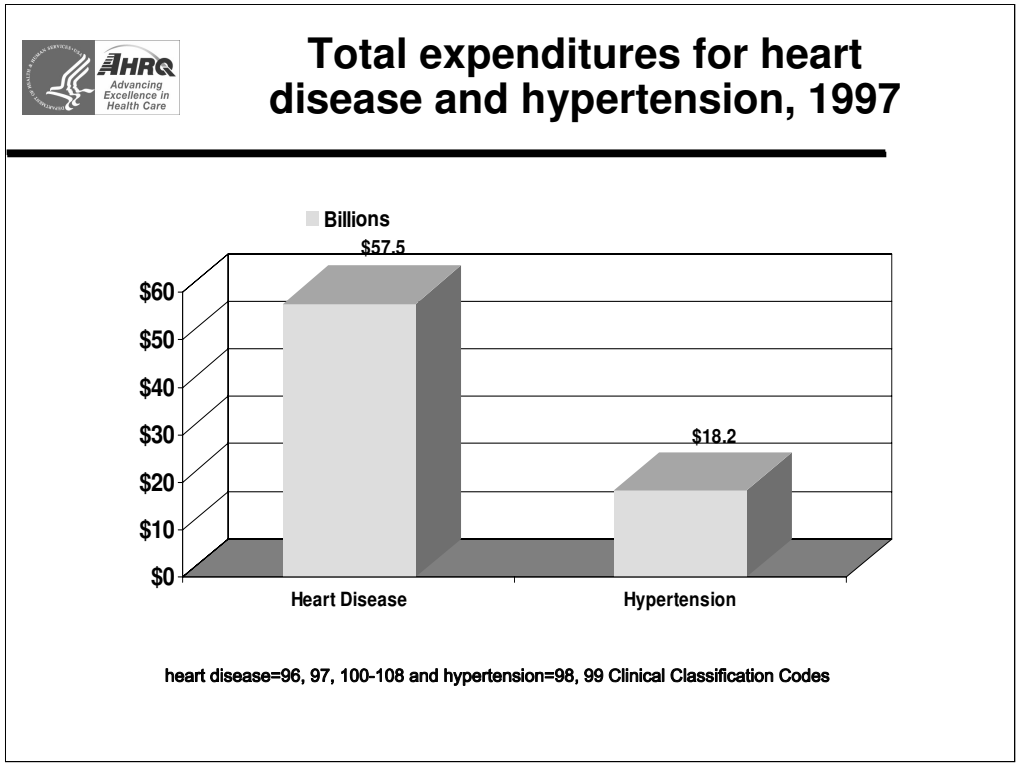
Each file has a *key* variable that uniquely identifies a row. For example, the full-year consolidated data file is a *person-level* file; i.e., each row has data for one person and different rows represent different persons. The key variable on this file is DUPERSID.

The medical conditions file is a *person-condition-level* file. The key variable CONDIDX is composed of DUPERSID plus a condition number. Each CONDIDX (each row on the condition file) uniquely identifies a condition for a person.

Note that the condition number (CONDN) is not sequential and cannot be used to count the number of conditions for a person.

ID variables are also important for merging different files. Merging files allows you to make connections between different files, e.g., between the medical conditions file and the hospital stays event file--or between the conditions file and the PMED file.

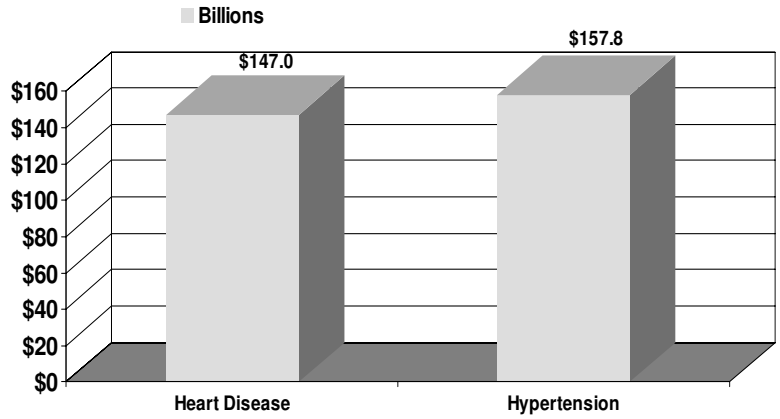
EXAMPLES OF GRAPHS USING MEDICAL CONDITIONS DATA



Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension

Total expenditures for people with heart disease and hypertension, 1997

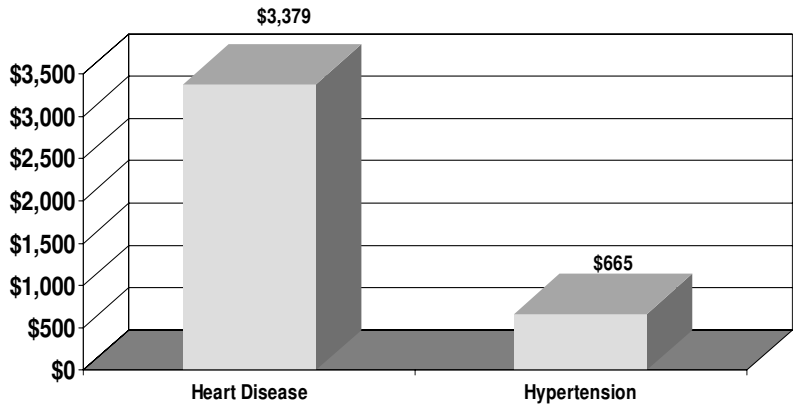


heart disease=96, 97, 100-108 and hypertension=98, 99 Clinical Classification Codes

Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension

Mean expenditures for heart disease and hypertension, 1997

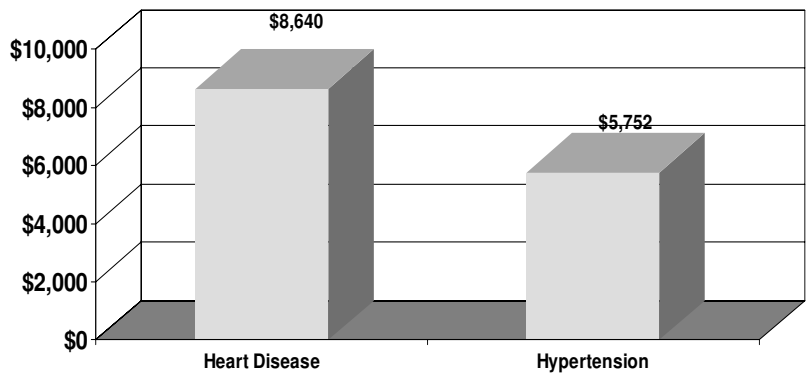


heart disease=96, 97, 100-108 and hypertension=98, 99 Clinical Classification Codes

Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension

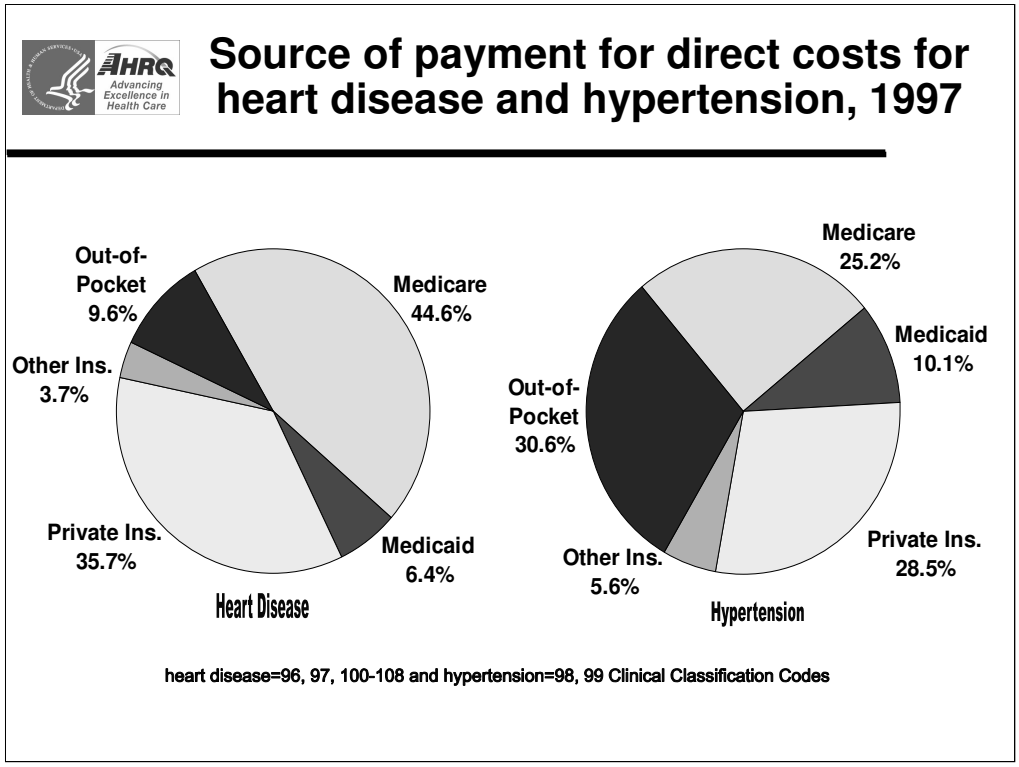
Mean expenditures for persons with heart disease and hypertension, 1997



heart disease=96, 97, 100-108 and hypertension=98, 99 Clinical Classification Codes

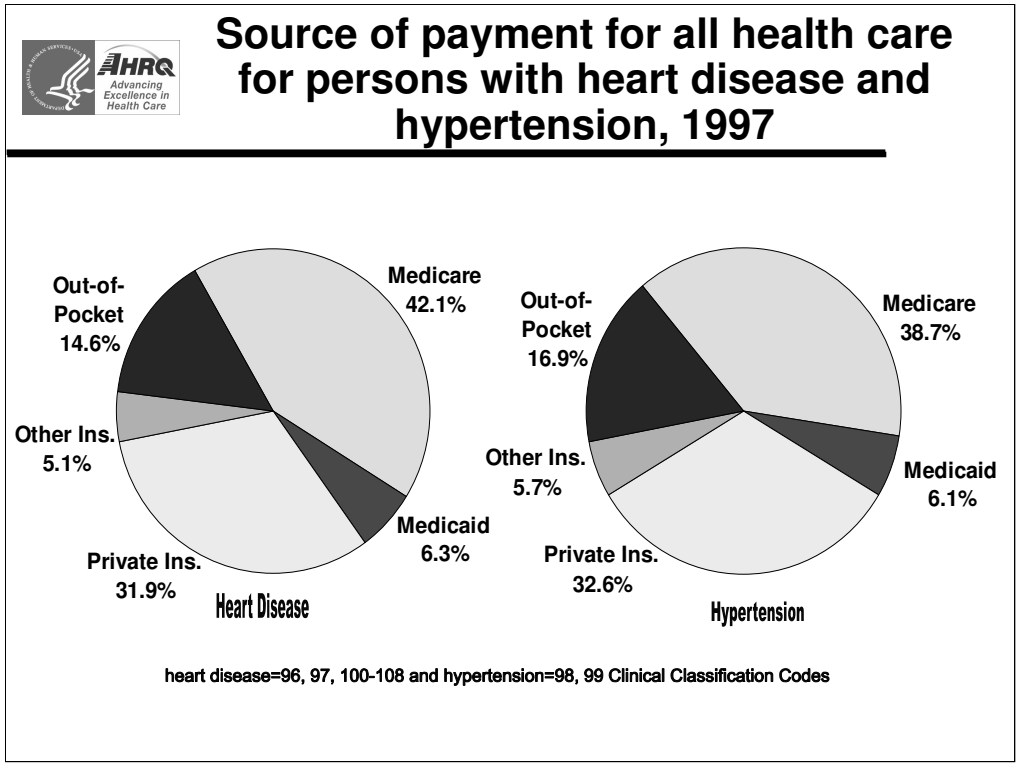
Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension



Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension



Heart Disease consists of the following Clinical Classification Codes:
96 – Heart valve disorders
97 – Peri-, endo-, and myocarditis, cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease)
100 – Acute myocardial infarction
101 – Coronary atherosclerosis and other heart disease
102- Nonspecific chest pain
103 – Pulmonary heart disease
104 – Other and ill-defined heart disease
105 – Conduction disorders
106 – Cardiac dysrhythmias
107 – Cardiac arrest and ventricular fibrillation
108 – Congestive heart failure, nonhypertensive

Hypertension consists of the following Clinical Classification Codes:
98 – Essential hypertension
99 – Hypertension with complications and secondary hypertension